DTIC FILE COPY

AD-A233 003

(2)

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

DTIC ELECTE MAR 14 1991 S B D

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 91 0157
6a. NAME OF PERFORMING ORGANIZATION: Center for Research in Math and Science Education	6b. OFFICE SYMBOL (If applicable)
7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research	
6c. ADDRESS (City, State and ZIP Code): San Diego State University, San Diego, CA 92182-0315	7b. ADDRESS (City, State and ZIP Code): Building 410, Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-89-0107	
8c. ADDRESS (City, State and ZIP Code): Building 410, Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2313	A4	

11. TITLE (Include Security Classification): Schema-based Theories of Problem Solving

12. PERSONAL AUTHOR(S): Stephen K. Reed

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Final Report	FROM 11/1/87 TO 12/31/90	1991 January 17	39

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Schema, examples, procedures, analogs, categorization, word problems

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The objective of this project was to formulate a schema-based model of problem solving to represent the construction of equations for solving algebra word problems. The first section shows how this research can be related to five characteristics of schema theories: abstraction of relevant properties, instantiation of values, predictions about expected information, induction from examples, and hierarchical organization according to different levels of specificity. The second section discusses the selection of analogous problems to determine whether students are able to judge which of two analogous problems will be more helpful. The third section discusses transfer to similar problems in which subjects must modify a solution to solve a problem that has the same story content. The fourth section discusses transfer to isomorphic problems in which subjects must overcome the different story content to detect the identity of two solutions. The fifth section discusses the variables that influence the successful categorization of problems according to common solution principles. (OVER)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. Alfred R. Fregly	22b. TELEPHONE NUMBER (Include Area Code): (202) 767-5021
22c. OFFICE SYMBOL: NL	

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

91 3 11 116

A concluding section compares the schema approach with a search approach that uses general heuristics to find a solution.

Summary

The objective of this project was to formulate a schema-based model of problem solving to represent the construction of equations for solving algebra word problems. The first section shows how this research can be related to five characteristics of schema theories: abstraction of relevant properties, instantiation of values, predictions about expected information, induction from examples, and hierarchical organization according to different levels of specificity. The second section discusses the selection of analogous problems to determine whether students are able to judge which of two analogous problems will be more helpful. The third section discusses transfer to similar problems in which subjects must modify a solution to solve a problem that has the same story content. The fourth section discusses transfer to isomorphic problems in which subjects must overcome the different story content to detect the identity of two solutions. The fifth section discusses the variables that influence the successful categorization of problems according to common solution principles. A concluding section compares the schema approach with a search approach that uses general heuristics to find a solution.

Introduction

The objective of this 3-year project was to develop a schema-based theory of problem solving that could serve as a basis for improving students' ability to solve algebra word problems. It began as a continuation of the PI's work on the use of analogous solutions (Reed, Dempster, & Ettinger, 1985; Reed, 1987). Our paradigm typically required that students construct equations to represent word problems. Because most college students have a lot of difficulty with this task, they are very dependent on examples. We were therefore interested in how the relation between the example and the test problem influences their performance on the test problem.

Our initial study on this topic explored transfer to problems that were either equivalent or similar to the examples (Reed, Dempster, & Ettinger, 1985). Transfer was reasonably good to equivalent problems but was not very good to similar problems that differed slightly from the example. Students were usually unable to adjust to changes between the example and test problem and would often try to use the same equation to solve the test problem without any modification. They were somewhat better at applying isomorphic solutions than at applying similar solutions (Reed, 1987) but even here there was room for improvement. Much of our research in this project was therefore motivated by attempts to find effective instructional procedures that would make examples more effective.

Research needs a unifying framework and I have found it convenient to think about my work within the context of schema theory. The first section of this report therefore discusses the major assumptions of schema theory and shows how they are relevant to my research on algebra word problems. The second section discusses the selection of analogous problems to determine whether students are able to judge which of two analogous problems will be more helpful. The third section discusses transfer to similar problems in which subjects must modify a solution to solve a problem that has the same story content. The fourth section discusses transfer to isomorphic problems in which subjects must overcome the different story content to detect the identity of two solutions. The fifth section discusses the variables that influence the successful categorization of problems according to common solution principles.

This report summarizes the results of a large number of experiments and therefore deletes many details. It focuses on those experiments that have been included in articles and manuscripts and does not include all of the experiments described in my two previous reports. Interested readers can consult these reports for more detail, or the articles and manuscripts listed at the end of this report for all the details.

Schema Theory

Characteristics of Schema Theories

There are many introductions to schema theory but I particularly like a chapter by Perry Thorndyke (1984). Thorndyke claimed that while schema theories are difficult to test and therefore underdeveloped as a descriptive theory, they provide principles for formulating a prescriptive theory. Prescriptive theories are particularly relevant for studying problem solving because attempted solutions can be scored as either correct or incorrect, providing a measure of the effectiveness of instruction. Can we therefore use the principles of schema theory as an organizational framework for comparing alternative approaches to increasing transfer?

Before attempting to make some comparisons, I will define and list the properties of a schema, based on Thorndyke's paper. Thorndyke (1984, p. 167) defines a schema as a cluster of knowledge representing a particular generic procedure, object, percept, event, sequence of events, or social situation. This cluster provides a skeleton structure for a concept that can be "instantiated", or filled out, with the detailed properties of the particular instance being represented. For example, a schema for the American Psychological Association annual meetings would contain the standard properties of a scientific conference such as its location, date, attenders, session types, and the length of presentations.

Thorndyke listed five characteristics of schema models that are generally shared by theorists who propose these models. The five characteristics are abstraction, instantiation, prediction, induction, and hierarchical organization.

A schema represents a prototypical *abstraction* of the concept it represents, listing those properties that define a typical instance. These properties provide the skeleton structure for the concept. The properties are represented as variables that can be filled out or *instantiated* by the particular instances that fit the schema. If the incoming information is incomplete, the schema may allow *predictions* about expected information and guide the interpretation of incoming information to match these expectations.

Schemata are formed by *induction* from experience with various instances of the general concept. This presumably occurs through a process of successive refinement in which expected properties become more clearly defined. And finally, schemata are *hierarchically organized* according to different levels of specificity. For example, expectations about the format of a particular convention are based on what typically occurs at professional conventions, in general.

The assumptions of schema theories provide a framework for discussing the prerequisites for the successful transfer of knowledge. In the next section I suggest how knowledge about algebra word problems fits this framework.

Algebra Word Problems

Algebra word problems require that students convert a word problem into an algebraic equation in order to solve for an unknown variable. The following is a typical problem:

A car travelling at a speed of 30 mph left a certain place at 10:00 a.m. At 11:30 a.m., another car departed from the same place at 40 mph and travelled the same route. In how many hours will the second car overtake the first car? (Problem 1)

Let's now consider how the five characteristics of schema theory - abstraction, instantiation, prediction, induction, and hierarchical organization - are relevant for solving this problem (Reed, 1991).

1. Abstraction. Abstraction requires learning those properties of a problem that are needed to solve it. The above example is an overtake problem in which the distance travelled by one car equals the distance travelled by another car. However, the distances in these problems are seldom stated directly, but must be represented as the product of rate and time. The correct equation is

$$\text{Rate1} \times \text{Time1} = \text{Rate2} \times \text{Time2} \qquad (1)$$

The rate of travel multiplied by the time of travel for one vehicle equals the rate of travel multiplied by the time of travel for the second vehicle.

In order to construct a correct equation students must learn that only the two rates and the two times are relevant for solving the problem. This knowledge is particularly important when there are irrelevant quantities in the problem, as in the following example (Krutetski, 1976, p. 110).

A train departed from city A to city B at a speed of 48 km per hour. Two hours later a second train followed it at 56 km per hour. At what distance from the starting point will the second train overtake the first, if the distance between the cities is 1200 km, and there are twice as many cars on the first train as on the second (Problem 2).

Krutetski found that capable students were not impeded by unnecessary data, whereas incapable students were often confused when such data were introduced into the text of even the easiest problems. All quantities were perceived by the incapable students as being equally important. We would therefore hope that students who were shown a solution to Problem 1 would learn that the distance between the cities and the number of cars on the train were irrelevant values in Problem 2.

I am unaware of recent research that has investigated irrelevant information in algebra word problems, perhaps because the problems are so difficult even without irrelevant information. However, information that is relevant in one problem may be irrelevant in another problem so we need to consider this issue in the study of transfer. Successful transfer depends on determining

what part of the solution must be mapped from a source problem to a target problem (Gentner, 1983).

2. Instantiation of Values. The solution of the overtake problem (Problem 1) requires not only that students identify the relevant quantities but that they substitute the correct quantities in the 'slots' of the equation. An equation provides a frame-like structure, but many problems are difficult even when students are given the correct equation because they don't know how to enter the correct values (Ross, 1989, Reed & Ettinger, 1987).

For example, several steps are required to represent the time travelled by the earlier car in Problem 1. First, the problem solver must calculate the time difference between the two cars by finding the difference between 10:00 a.m. and 11:30 a.m. Second, if the time travelled by the later car is represented by the variable t, then the time travelled by the earlier car is represented by the variable t + 1.5 hours. Students must know that they should add, rather than subtract, 1.5 hours. And finally, students must be careful to pair each of the times with the corresponding rate of travel. Students who had learned to carry out these operations in solving Problem 1 would hopefully be able to apply them in solving Problem 2.

3. Predictions Schemas should make predictions about expected information and guide the interpretation of this information. I will distinguish between two kinds of expected information. One kind concerns the default values that go into the slots of a schema. For example, if most college courses are for 3 credits, then one might reasonably infer that a course is worth 3 credits in the absence of information. The expectation of a particular value is unreasonable for most word problems, however, because the quantities vary from problem to problem. As a contrast, a frame-based system for diagnosing diseases can make use of expected values by determining which values of a patient's medical profile fall outside the normal range of values (Aikens, 1983).

A second kind of expectation requires knowledge of what information is required in order to solve a particular problem. In addition to studying students' ability to recognize irrelevant information, Krutetski (1976) studied their ability to recognize when they had insufficient information. As was found for the identification of irrelevant data, only the most capable students were able to identify what information was missing from a problem.

We would probably have many irate students if we asked them to solve problems with missing information. But, as was the case for irrelevant information, an analogous problem may have missing information. We have referred to these problems as less inclusive than the test problem and to problems that have excess information as more inclusive than the test problem (Reed, Ackinclose, & Voss, 1990).

4. Induction. Induction is concerned with how schemata are formed from experience with instances of the concept. Hinsley, Hayes, and Simon (1977) proposed that if schemata are important for solving algebra word problems, then problem categories should guide students' solutions. They tested this hypothesis by asking high school and college students to sort by 'problem type' 76 word problems taken from an algebra textbook. Students sorted the problems into 16 - 18 categories and there was considerable agreement among the students regarding the identity of these categories. (such as distance, interest, area, mixture, and work problems). The experimenters suggested that information about the problem categories, including relevant equations and diagrams, is useful for formulating solutions.

A limitation of this knowledge is that each category contains a variety of problems, each requiring a slightly different solution. Mayer (1981) refers to these variations as *templates* and developed a taxonomy of problem types from a collection of over 1,000 word problems taken from 10 textbooks. Thus the overtake problem in Mayer's taxonomy is only one of 13 templates composing motion problems and these do not include river current problems. Other examples include one car travelling at different rates on a round trip and two cars travelling toward each other to meet. Unfortunately, rather small changes in a problem can greatly reduce the effectiveness of an example (Reed, Dempster, & Ettinger, 1985; Reed & Ettinger, 1987). Thus one of the challenges for a prescriptive theory is to formulate how knowledge should be organized to allow students to effectively solve the variations of problems that exist in each of the categories identified by Hinsley et al. (1977).

5. Hierarchical Organization. Another challenge for a prescriptive theory is to specify how to organize knowledge at more general levels. Hinsley's subjects classified problems by using similarities in story content, and could often make their classifications after hearing only a few words. With increasing expertise, people are able to classify problems according to the similarity of solutions rather than the similarity of story content (Chi, Glaser, & Rees, 1982; Schoenfeld & Hermann, 1982; Silver, 1981).

The FERMI system developed by Larkin, Reif, Carbonell, & Guglietta (1988) to solve physics problems is a good example of a system that exploits hierarchical organization through its knowledge of general principles and methods that apply to a large variety of domains. For example, one general method is decomposition which decomposes a complex problem into simpler ones associated with its components. A combination function specifies how a desired quantity can be found from the quantities associated with the individual components. An advantage of a hierarchical organization is that general knowledge needs to be encoded only once and then can be

used repeatedly on specific problems that share common solutions. I will return to this issue when I discuss isomorphic problems later in this chapter.

Summary. In conclusion, algebra word problems fit rather nicely into a schema interpretation. Students typically categorize problems and use equations associated with problem categories as a basis for solving the problems. The equations consist of concepts such as distance, rate, and time that represent the features of the different categories. These concepts are combined to form an equation and are replaced by instantiated values (numbers or variables) when the equation is applied to solve a particular problem.

The initially learned categories are based on story content (Hinsley et al., 1977) and could be considered as basic categories in Rosch's taxonomy (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). The subordinate categories would then be the templates in Mayer's (1981) taxonomy -- the different variations in distance problems, mixture problems, etc. The superordinate categories would consist of isomorphic problems that share a common solution but have different story content. Acquiring expertise requires developing skills in both directions from the basic level. Students must learn how to modify an equation of a typical problem to create variations in order solve similar problems within a category. They must also learn how to recognize isomorphic problems in order to apply the same general method across different categories.

Selecting Analoguos Problems

In a recent paper on analogical problem solving, Holyoak and Koh (1987) identified four basic steps in transferring knowledge from a source domain to a target domain: (1) constructing mental representations of the source and the target; (2) selecting the source as a potentially relevant analogue to the target, (3) mapping the components of the source and target; and (4) extending the mapping to generate a solution to the target. They stated that the second step, selecting a source analog, is perhaps the least understood of the four steps.

The objective of our study (Reed, Ackinclose, & Voss 1990) was to identify variables that influence the selection of analogous problems and to determine whether students would select effective solutions. In the first experiment students had to choose between two problems that belonged to the same category as the test problem. One problem was less inclusive than the test problem and the other problem was more inclusive than the test problem. The less inclusive problem had missing information and the more inclusive problem had excess information. In the second experiment students had to choose between a problem that was less inclusive than the test problem and a problem that was isomorphic to the test problem.

The same pattern of results occurred in both experiments: students selected problems on the basis of perceived similarity. They did not show a significant preference for the more inclusive problems in the first experiment or the isomorphic problems in the second experiment although both sets of solutions were significantly more effective than solutions to the less inclusive problems. The results therefore reveal a discrepancy between the variable that determines the selection of solutions (similarity) and the variable that determines the usefulness of solutions (inclusiveness).

The purpose of the third experiment was to determine whether either mathematical experience or the opportunity to study the solutions of analogous problems would increase students' ability to select good analogies. The subjects in the first two experiments were tested in college algebra classes and therefore had similar preparation in mathematics. In contrast, the subjects in the third experiment were participants in the psychology subject pool and therefore had a more varied background in college mathematics courses. They were classified according to three levels of experience, depending on whether they (1) had not taken college algebra, (2) had taken college algebra, or (3) had taken calculus. The second factor -- the opportunity to study the solutions -- was varied by allowing subjects to study the solutions to half of the problem sets before they made their selections.

The results showed that neither more experience nor seeing the solutions increased the selection of more-inclusive solutions. Experience also did not influence the selection of isomorphic solutions, although seeing the solutions did significantly increase the selection of isomorphs.

The failure to find an effect of mathematical experience on selecting solutions may have been caused by an insufficient range in experience. In the fourth experiment, we also included a group of undergraduates who were majoring in mathematics and planned to teach mathematics at a junior high or high school. They were all enrolled in an upper-division mathematics course, Basic Mathematic Concepts, and had previously taken an average of 6 mathematics courses.

The analysis for the similar problems again revealed that experience did not significantly influence subjects' preferences. The more-inclusive solution was selected on 38% of the occasions for students who had not taken college algebra, 39% for students who had taken college algebra, and 42% for students who were mathematics majors. In contrast, mathematical experience did have a significant effect on the selection of isomorphic problems. The isomorphic solution was selected on 37% of the occasions for students who had not taken college algebra, 28% of the occasions for students who had taken college algebra, and 50% of the occasions for students majoring in mathematics.

The finding that students majoring in mathematics did better in selecting isomorphic problems is consistent with the results of Experiment 3 that showing students the solutions of the

analogous problems improved the selection of isomorphs. It is noteworthy that although both of these variables influenced the selection of isomorphs, neither of these variables influenced the selection of similar problems. It is apparently difficult to identify a good analogy when neither of the two analogous problems has a solution that is identical to the test problem. The development of training procedures to help students select good analogies may therefore be beneficial for improving performance on test problems.

Transfer to Similar Problems

Examples plus Procedures

The difficulty with too little information is that people have to generate the missing information and incorporate it into the equation. A possible remedy for generating missing or altered quantities is to build instruction around procedures that can generate quantities that are in the test problem but not in the example. Procedures can then be attached to slots in a schema for the purpose of generating values to fill the slot (Bobrow & Winograd, 1977; Larkin et al., 1988). For instance, in a work problem the two workers may work together either for the same number of hours or for a different number of hours. Students who learn how to solve the former problem need to know how to modify the solution when one worker labors for more hours than the other worker.

To assess the effectiveness of combining an example with a set of procedures, Reed and Bolstad (1991, Experiment 1) compared three different methods of instruction. In the Example condition students received a detailed solution to the following problem:

Ann can type a manuscript in 10 hours and Florence can type it in 5 hours. How long will it take them if they both work together? (Problem 3)

In the Procedures condition students were given the correct equation along with a set of rules that specified how to enter values into the equation. The rules for representing time were:

1. Time refers to the amount of time each worker contributes to the task. If this value is stated in the problem, enter it into the equation. For example, if one person works for 5 hours, enter 5 hours into the equation for that worker.

2. Time is often the unknown variable in these problems. Be sure to represent the correct relative time among the workers if they do not work for the same time. If one worker works 3 hours more than the value ($\underline{h}$) you are trying to find, enter $\underline{h}$ + 3 for that worker.

In the Example & Procedures condition students received both the example solution and the set of rules. After studying the instructional material for 5 minutes, all students attempted to construct equations for 8 test problems that differed from 0 to 3 transformations from the example.

The transformations modified either *rate* (expressed relationally for the two workers), *time* (one worker labored longer), or *tasks* (part of the task had already been completed). Each of these is illustrated in the test problem that differed by 3 transformations from the example:

John can sort a stack of mail in 6 hours and Paul is twice as fast. They both sort 1/5 of the stack before their break. How long will it take John to sort the remainder if he and Paul work together, but Paul works 1 hour longer? (Problem 4)

The 4 transformation levels and 3 instructional methods enabled us to evaluate the predictions of a model for each of these 12 conditions. Because both the example and procedures provide students with the basic equation for solving these problems, we assume that the probability of generating a correct equation is equal to the probability of correctly generating the values for the 5 quantities in the equation: the rate and time of work for each of the two workers and the number of tasks to complete. Students can generate these values by either matching the information provided in the example (m), following the rules provided in the procedures (r), or using their general knowledge about the problems (g). The parameters m, r, and g specify the probability of generating a correct value from each of these sources of knowledge.

Let's first consider the predictions for the instructional group who receives the example and the procedures. When the test problem is equivalent to the example, a student can generate all 5 values by using the matching operation. The probability of generating a correct equation is therefore m^5 -- the probability that the student correctly applies the matching operation to each of the values in the example. When the test problem differs by 1 transformation the probability of a correct equation is m^4r. In this case the student can match 4 of the quantities but must use the procedures to generate the transformed value. Following the same logic, the probability of correctly generating an equation should be m^3r^2 for 2 transformations and m^2r^3 for 3 transformations. Assuming it is easier to match values in the example than follow procedures ($m > r$), the model predicts a decline in performance as the number of transformations increases.

When students have only the example, they must rely on their general knowledge to generate the transformed quantities. The probability of constructing a correct equation should therefore be m^5 for 0 transformations, m^4g for 1 transformation, m^3g^2 for 2 transformations, and m^2g^3 for 3 transformations. The generalization gradient should be steeper for the Example group than for the Example & Procedures group if the rules increase the probability of correctly generating the transformed values ($r > g$).

When students have only the rules, there should not be a generalization gradient. In this case, the probability of constructing a correct equation should be r^5 -- the probability of correctly applying a rule to generate each of the 5 values.

Table 1 shows the observed and predicted values for each of the 12 instructional situations, based on parameter estimates of .96 for correctly matching information in the example, .65 for following a procedure, and .45 for using general knowledge. The model accounted for 94% of the variance in the data, as determined by the square of the multiple regression coefficient. The first column illustrates the steep generalization gradient that occurs as the test problems become more dissimilar from the example. The fourth column illustrates that the rules were ineffective when presented alone and need to be clarified or supplemented with additional material. For instance, a set of procedures for teaching people how to operate a device can be facilitated with the addition of functional, structural, or diagrammatic information that enables better understanding and integration of the rules (Keiras & Bovair, 1984, Smith & Goodman, 1984, Viscuso & Spoehr, 1986).

The combined effect of the example and procedures was disappointing from an instructional perspective. The data in Table 1 show that the generalization gradient for the Example & Procedures group was not quite as steep as for the Example group, but was steeper than predicted by the model. Furthermore, the overall performance of these two groups was not significantly different, although both differed significantly from the Procedures group.

The results of this experiment demonstrated that a potentially effective way of organizing knowledge was unimpressive in application. Of course, our application may have been faulty and we attempted (unsuccessfully) to improve the procedures in a subsequent experiment. We also included an instructional condition in which students received a second example. This condition, which was more successful, is discussed in the next section.

Multiple Examples

A possible remedy for poor generalization among the variations of a problem is to present many examples to show the variations. Sweller and Cooper (1985) proposed that students need to be shown a wide range of worked examples in order to become proficient in solving problems. They based their conclusion on the finding that (1) students did better when studying worked examples of algebra manipulation problems than when attempting to solve the examples, and (2) the superiority of worked examples was specific to test problems identical in structure to the examples.

Reed and Bolstad (1991, Experiment 2) compared the multiple-example method with the example-plus-procedures method on the variations of the work problem discussed in the previous section. Students in the two-example condition received solutions to the simple example (problem 3) used in the previous experiment and to a complex example that was equivalent to the 3-transformation problem (Problem 4). Other groups received either a single example, a set of

Table 1

Observed and Predicted Values for Experiment 1

Transformations	Groups								
	Example			Procedures			Example & Procedures		
	Observed	Predicted	Model	Observed	Predicted	Model	Observed	Predicted	Model
0	82	81	m^5	19	12	r^5	82	81	m^5
1	36	37	m^4g	19	12	r^5	42	55	m^4r
2	18	17	m^3g^2	17	12	r^5	30	37	m^3r^2
3	0	7	m^2g^3	5	12	r^5	14	25	m^2r^3

Note. The predictions are based on parameter estimates of m = .96 (the probability of correctly matching the example), r = .65 (the probability of correctly applying a rule), and g = .45 (the probability of correctly applying general knowledge).

procedures, or an example and procedures. The percentage of correct equations on 8 test problems for the different instructional conditions was 7% for the procedures, 32% for the complex example, 38% for the simple example, 45% for the complex example and procedures, 47% for the simple example and procedures, and 65% for the simple and complex examples. The group which received only the procedures performed significantly worse than all other groups and the group which received two examples performed significantly better than all other groups. Other differences were not significant.

The performance of the two-example group across the eight test problems suggested that the complexity of the values in the test problem was a more limiting factor on performance than the number of examples required to solve the test problem. A complex value occurs when either rate is expressed as a relation between workers, time is unequal across workers, or part of the task is completed. The percentage of correct equations was 90% when the test problem contained no complex values, 70% for 1 complex value, 66% for 2 complex values, and 55% for 3 complex values. A problem with 0 complex values is equivalent to the simple example and a problem with 3 complex values is equivalent to the complex example. The test problems between these two extremes required combining information from both examples.

The advantage of being able to combine information across two examples is that it eliminates the necessity of having to provide a training example for every possible test problem. The 8 test problems in the Reed & Bolstad experiment represented the 8 possible combinations created from 3 attributes (rate, time, tasks) with 2 values (simple, complex). In order to maximize transfer from 2 examples it is necessary to establish the independence of an attribute across changes in the other attributes. If students learn how to represent time when one worker labors more than the other, then they should ideally be able to transfer this knowledge to other problems when the other attributes (such as rate or tasks) change values. It was this assumption of independence that allowed Reed and Bolstad to represent 8 test problems by 2 examples.

<u>Summary</u>

Algebra word problems have the potential advantage of being recognized as members of well-defined categories (Hinsley et al., 1977). The categories provide a starting point for formulating a solution to problems in the category. Solving a work problem, for example, usually requires representing the amount of work accomplished by a worker as a product of her rate of work and the amount of time she works. However, there can be many variations of problems within a category, requiring that students learn how to modify their solutions to fit these various templates (Mayer, 1981).

The potential advantage of classifying a problem as a mixture, motion, distance, work or some other categorical problem is limited by students' inability to adjust to small changes in solutions. In fact, the evidence suggests that isomorphic solutions from a different category provide more useful information than similar solutions that belong to the same category (Reed, 1987). The limited usefulness of similar solutions can be improved by using solutions that are more inclusive than the test problem, but students do not seem to be sensitive to this dimension when choosing an analogous problem (Reed et al., 1990).

Another approach, based on procedural attachments, is to supplement a similar example with a set of procedures that specify how to enter values into the equation. Our attempt to improve transfer to similar problems by giving students an example and a set of procedures has thus far been less successful than giving students two examples which span the set of test problems (Reed & Bolstad, 1991) The success of the latter approach is consistent with the multiple-examples approach advocated by others (Catrambone & Holyoak, 1990, Sweller & Cooper, 1985) and demonstrates that students can selectively choose the appropriate quantities from each problem. Our finding that we needed only 2 examples to obtain good transfer to 8 test problems is an encouraging verification of this approach.

Transfer to Isomorphic Problems

The ability to modify a solution to solve a similar problem is an important skill in becoming an efficient problem solver. Another important skill is to be able to recognize and apply a solution from an isomorphic problem. Because an isomorphic solution has the same structure as the test problem, the problem solver has to appropriately map the concepts in the solution onto the concepts in the test problem, rather than modify the solution.

According to a theory proposed by Gick & Holyoak (1983), recognition of isomorphic problems is facilitated by schema abstraction in which corresponding concepts in the two solutions are recognized as specific instances of a more general concept. The next section discusses increasing transfer by representing problems at a more general level.

Schema Abstraction

Gick and Holyoak's (1983) theory of schema abstraction was the outcome of a series of attempts to improve students' noticing an analogous solution to Duncker's radiation problem. The problem requires using radiation to destroy a tumor without harming the healthy tissue that surrounds it. A convergence solution involves dividing the rays so they will have a high intensity only when they converge on the tumor.

Before attempting to solve the radiation problem some students read a story about a general who was able to capture a fortress by dividing his army along different roads so it could converge on the fortress. Because of mines on the roads, the army could not attack along a single road. The solution to the military problem was quite helpful when subjects were instructed to make use of the story when solving the radiation problem, but most subjects failed to notice the analogy between the two problems when not given a hint (Gick & Holyoak, 1980).

Noticing an analogy between isomorphic problems is limited by different concepts, such as army and fortress in the military problem and rays and tumor in the radiation problem. The similarity of the two solutions becomes apparent when the concepts are described at a more general level such as using force to overcome a central target. Gick and Holyoak (1983) found that their students were likely to form this more general schema if they read and compared two analogous stories before trying to solve the radiation problem. For example, some students read the military story and a story about forming a circle around an oil fire in order to use many small hoses to spray foam on the fire. Students who described the relation between these two stories were more likely to think of the convergence solution to the radiation problem than students who read only a single analogous story.

The benefit of creating a more general schema was demonstrated by Catrambone and Holyoak (1989) who used more-directed comparison instructions to explicitly require the creation of superordinate concepts. All subjects read the military and fire stories and wrote summaries of them. But some subjects received the following two statements:

1. The fortress is difficult to capture because a large army of soldiers can not attack it from one direction.
2. The fire is difficult to put out because a large amount of water can not be thrown at it from one direction.

They were then instructed to write a third sentence in which the pairs of concepts *fortress* and *fire*, *army* and *water*, and *attack* and *thrown at* are replaced by a more general term. They were then shown the more general statement:

3. A target is difficult to overcome because a large force can not be aimed at it from one direction.

Subjects who received the more-directed comparison instructions and applied it to a third analogue were more successful in producing the convergence solution than students who received the less-directed set of comparison instructions.

<u>Abstraction of Word problems.</u>

Although these results are impressive, it is unclear whether they could be duplicated for complex problems such as algebra word problems. Research on students' ability to categorize problems according to the formal procedures required to solve them has shown that correct classification requires considerable expertise (Chi, Glaser, & Rees, 1982) or training (Schoenfeld & Hermann, 1982). Can detailed comparisons of isomorphic problems therefore result in the creation of abstract solution procedures for word problems?

Dellarosa (1985) was partially successful in using analogical comparisons to help students identify which word problems shared a common solution. Students were trained on problems that had three different story contents (vat, travel, or interest) and three different solution structures. She found that students who compared quantities and relations in one problem to quantities and relations in an isomorphic problem did significantly better in classifying problems according to the common solutions than students who answered questions about individual problems. In a second experiment, Dellarosa found that students who did analogical comparisons also were more accurate in matching word problems to equations than students who answered questions about individual problems. However, the analogical comparisons were not successful in helping students use the equations to solve problems.

Reed (1989) used three isomorphic mixture problems and three isomorphic distance problems to determine whether constructing an analogical mapping between two isomorphs would help students solve a third isomorph. In the mixture problems (the nurse, grocer, and Mr. Smith problems in Table 2), two quantities are added together to make a combined quantity. These problems are solved the same way but belong to different categories in Mayer's (1981) taxonomy. In the distance problems, two distances are added together to equal the total distance but the two objects are either travelling toward each other, travelling away form each other, or succeeding each other. These problems belong to the same category in Mayer's taxonomy, but represent different templates and are characterized by different spatial relations (converge, diverge, succession) between the two moving objects.

Subjects in my experiment attempted to construct equations for the series of 3 mixture problems and 3 distance problems which were counterbalanced for order of presentation. Providing an analogous solution helped students construct equations. The percentage of correct solutions increased from approximately 10% for the the first problem in a series to 50% for the second problem, which was accompanied by the solution to the first problem. However, a detailed comparison of the first two problems through mapping concepts did not increase the solution rate for the third problem. Furthermore, success on the third problem was uninfluenced by whether

Table 2

Isomorphic Mixture and Distance Problems (Reed, 1989)

Number	Problem
	Mixture
1 (Wet)	A nurse mixes a 6% boric acid solution with a 12% boric acid solution. How many pints of each are needed to make 4.5 pints of an 8% boric acid solution?
1E (Wet)	A chemist mixes a 20% alcohol solution with a 30% alcohol solution. How many pints of each are needed to make 10 pints of a 22% alcohol solution?
2 (Dry)	A grocer mixes peanuts worth $1.65 a pound and almonds worth $2.10 a pound. How many pounds of each are needed to make 30 pounds of a mixture worth $1.83 a pound?
2E (Dry)	A candy dealer mixes peppermint worth $0.75 a pound and butterscotch worth $0.90 a pound. How many pounds of each are needed to make 9 pounds of a mixture worth $0.80 a pound?
3 (Interest)	Mr. Smith receives 5% interest from his checking account and 14% interest from treasury bonds. How much money is in each account if he averages a 12% return on $4500?
3E (Interest)	Mr. Roberts receives 7% interest from stock dividends and 11% interest from his IRA account. How much money is in each account if he averages an 8% return on $8000.
	Distance
1 (Convergence)	Mary and Sue live 50 miles apart. They decide to ride their bicycles toward each other to meet for a picnic. Mary rides at 10 mph. Sue leaves 2 hours after Mary and rides at 8 mph. How long will Sue ride before they meet?
1E (Convergence)	Howard and Allan live 135 miles apart. They decide to drive toward each other, and Howard drives at 54 mph and Allan drives at 48 mph. How long will Allan drive before they meet if he leaves 1 hour after Howard?
2 (Succession)	Bill and Adam run a long-distance relay race for a total distance of 20 miles. Bill runs at 4 mph but runs 0.5 hours longer than Adam. Adam runs at 5 mph. How long does Adam run?
2E (Succession)	Sherry and Becky swim in a long-distance relay for a total distance of 12 miles. Sherry swims at 2 mph and Becky swims at 3 mph. How long does Becky swim if Sherry swims 1.5 hours more than Becky?
3 (Divergence)	A freight train leaves a station traveling at 45 mph. A passenger train leaves 3 hours later traveling at 60 mph in the opposite direction. How long will the passenger train have traveled when the two trains are 250 miles apart?
3E (Divergence)	A bus leaves a rest stop 2.5 hours before a truck leaves the same rest stop, traveling in the opposite direction. The bus travels at 55 mph and the truck travels at 50 mph. How long will the truck have traveled when the two vehicles are 400 miles apart?

students compared the first two problems or elaborated each individual problem. These findings failed to support a schema abstraction hypothesis.

One of the differences between the experimental paradigm used in this experiment and the paradigm used by Gick and Holyoak is that my students could refer to a specific analog (the solution of the second problem) as they worked on the third problem. It is possible that providing a solution to a specific analog may have discouraged students from using an abstract solution schema. This hypothesis was explored in a second experiment by not allowing subjects to refer back to the solution of the second problem as they worked on the third problem. Once again, however, the schema abstraction paradigm was unsuccessful.

In a third experiment I evaluated two other possible causes of the failure to create abstract solution schemas. One limitation is that students did not perform perfectly on the concept-matching task when comparing isomorphic problems. Because they did not receive feedback, some students did not know the correct mapping between all the lower-order relations. I tested this hypothesis in a third experiment by giving one group of students the correct answers on the concept-matching task. Another hypothesis is that practice on the concept-matching task does not help students learn the higher-order relations that are necessary to form an abstract solution schema. I tested this hypothesis by telling students the common principle for the first two problems in each series. Neither of these instructional variations was successful in improving performance on the third problem in the series.

Lack of Superordinate concepts.

I suggested several explanations of why abstraction did not occur, but I want to focus on the lack of superordinate concepts to describe a more general schema. As mentioned previously, superordinate concepts played an important role in the initial theoretical formulation of a convergence schema (Gick & Holyoak, 1983) and in subsequent instructional efforts to promote abstraction (Catrambone & Holyoak, 1989).

The creation of superordinate concepts has also been proposed by investigators working on artificial-intelligence approaches to analogical reasoning. Winston (1980) suggested that finding an analogy between two situations may require matching concepts at more general levels than those provided in the statement of the problem. However, he cautions that creating concepts that are too general will not sufficiently constrain the matches between two analogs. According to this hypothesis, abstraction requires creating concepts that are superordinate to the concepts in the isomorphic problems but are not so general that they do not sufficiently constrain the solution. A constraint on creating an abstract solution is that it may be difficult to find such concepts.

Because the concepts of distance, rate, and time do not differ in the distance problems, they do not have to be replaced by superordinates. Only the described action has to be generalized to replace the specific actions of travel toward in the convergence problem, travel successively in the succession problem, and travel away in the divergence problem. Although the superordinate concept travel generalizes each of these actions, it is too general to constrain the spatial relation between the two objects that are traveling.

The lack of superordinate concepts is even more apparent in the mixture problems. Although there are some identical concepts in specific pairs of problems (such as money in the interest and grocer problems), none of the specific concepts can be generalized under a superordinate concept that applies to all three problems. Thus volume is an important concept in the nurse problem, weight in the grocer problem, and interest rate in the interest problem.

Because it is difficult to create superordinate concepts for these problems it is necessary to consider alternative approaches for showing which problems have structural similarities. One option is to teach students at a more specific level than superordinate concepts. We will next consider an exemplar approach in which isomorphs are learned as specific examples, without schema abstraction.

Learning Common Principles

The following experiments on categorization investigate variables that influence students' ability to learn correct equations. A hierarchical model proposed by Reed (1987) claims that constructing an equation for a word problem requires the successful completion of three steps, beginning with the identification of the principle of the problem. For example, motion problems typically require that two distances either be equated, added, or subtracted. Learning the principle requires learning which problems belong in each of these three categories.

There is abundant evidence that the ability to identify the principle of a problem is an important component in acquiring expertise. As people become better problem solvers they group problems together on the basis of common principles and mathematical structure rather than on the basis of common objects and story context (Chi, Glaser, & Rees, 1982; Hardiman, Dufresne, & Mestre, 1989; Schoenfeld & Hermann, 1982; Silver, 1981). This finding from sorting studies has been confirmed by the collection of verbal protocols. The explanations generated by good students while studying worked-out examples of mechanics problems demonstrated that they could relate the solution to principles in the text (Chi, Bassock, Lewis, Reimann, & Glaser, 1989). Of particular relevance are the findings obtained by Berger and Wilde (1987) who developed a hierarchical task analysis of algebra word problems that is very similar to my analysis (Reed,

1987). Protocols from their subjects revealed that over half of the experienced students showed a "working-down" strategy that started at the top (principle) level, providing a relatively clear indication of the path to the goal. The novices never started at this level.

Importance of Templates

The purpose of our study (Reed & Sanchez, 1990) was to investigate how differences in the instructional examples influence the accuracy of classifying motion problems according to the arithmetic operation specified by the principle. Each problem contained two travelled distances that had to be equated, added, or subtracted. The examples were selected from the set of 12 problems shown in Table 3.

The four examples in each category can be further divided into two templates based on Mayer's (1981) taxonomy. According to Mayer, problems belong to the same template if they share the same story line and same list of propositions, regardless of the actual values assigned to each variable or which variable is assigned the unknown. The first and third problems in each category belong to one template (which we will refer to as Template 1) and the second and fourth problems belong to a second template (Template 2). The two examples within a template consist of one problem in which time is the unknown and one problem in which rate is the unknown.

Five of the six templates in Table 3 correspond to templates identified by Mayer. The equate problems consist of overtake and round-trip templates; the addition problems consist of closure and speed-change templates; and the subtraction problems include the same-direction template. The second pair of subtraction problems involves a speed change but the problems specify a difference in two distances rather than a total distance as in the addition problems.

A distinction between Template 1 and Template 2 problems is that all the Template 1 problems consist of a comparison between two objects or people and all the Template 2 problems consist of a single object or person travelling at two different speeds. Although this is a common feature across the three categories, there are also differences. For example, in the overtake problems (equate) the two people travel in the same direction, whereas in the closure problems (add) the two people travel in opposite directions.

A central concern of our study was to investigate the effect of describing templates at two different levels of generality. In Experiments 2 and 3 we contrast the descriptions of the specific problems shown in Table 3 with more general descriptions that emphasize the principles of equating, adding, and subtracting distances. However, we first demonstrate that templates are psychologically relevant by evaluating our belief that generalizations are conservative, and are mostly limited to recognizing common solutions among problems that belong to the same template.

Subjects in Experiment 1 were tested in a pretest-instruction-posttest design in which the *test* problems consisted of two problems representing each of the six templates. The *instructional* problems consisted of two examples to represent each of the three categories. Either the two examples were from the same template (such as two round-trip problems) or were from different templates (a round-trip and an overtake problem). Subjects could refer to the examples as they classified problems on the posttest.

Our first hypothesis was that people who received examples of only one template for each category would be significantly more accurate in classifying test problems that matched those templates than in classifying test problems that did not match those templates. Subjects in the same-template group improved from 31% correct classifications on the pretest to 59% correct classifications on the posttest for those problems that matched the instructional templates. In contrast, they did not improve in classifying test problems that did not match the instructional templates. These problems were classified with 37% accuracy on the pretest and 34% accuracy on the posttest. As predicted, the posttest performance was significantly higher on those problems that matched the templates presented as instructional examples.

Our second hypothesis was that students who received two different templates for each category would show significantly greater improvement than students who received two variations of the same template. Students in the different-template group improved from 45% to 57% correct classifications and students in the same-template group improved from 35% to 46% correct classifications. Performance was significantly higher on the posttest, but the group x test interaction was nonsignificant.

These results did not support the hypothesis that subjects who had two different templates would do significantly better on the posttest than subjects who had two variations of the same template. Although subjects in the same-template group did not improve on test problems that didn't match the instructional templates, they were helped by having two examples of the templates that were presented in the instruction. This partially compensated for the lack of templates for some of the problems.

Two variations of a problem can help people form more general descriptions of the problem (Gick & Holyoak, 1983), although this usually requires an active comparison of the two variations, either through instructions to create abstractions (Catrambone & Holyoak, 1989) or through the use of one problem to solve the other (Ross & Kennedy, 1990). It is therefore unclear whether students who received two examples of a template formed a more general description of that template or used the specific examples. One approach to studying this issue is to determine under what conditions giving students more general descriptions enhances their accuracy in classifying

problems. If students are capable of forming their own general descriptions from examples, then providing general descriptions should not enhance their performance. In Experiment 2 we investigated the usefulness of describing templates at two different levels of generality to compare the effectiveness of a general description with the effectiveness of specific examples.

Generality of Descriptions

The purpose of Experiment 2 was to compare two different levels of generality in describing the two templates that belong to each category. The *specific* descriptions are simply descriptions of the training problems that were printed in Table 3; the *general* descriptions are summaries of the problems that emphasize the principle of how the two distances are related. The general descriptions are at a low level of abstraction because they describe concrete situations such as going on a round trip or overtaking another person. But they omit the details of the problems such as whether a person or an object is being overtaken, the specific quantities, and the unknown variable. Table 4 shows the general description of each template.

We were particularly interested in whether the specific and general descriptions are sufficiently different that presenting both would be informative. Because we described the principle at the template level, rather than at a more abstract level, it may have been induced by subjects and therefore be redundant.

Evidence supports the claim that two levels of description are beneficial when one level is very abstract and general and the other is very concrete and specific. For instance, research on logical inferences has shown that presenting both general rules and specific examples is significantly better than presenting either alone (Cheng, Holyoak, Nisbett & Oliver, 1986). The general rules in the Cheng study were very abstract logical rules of the form *If p, then q.* Because abstract ideas may be difficult to comprehend (Bransford & Johnson, 1973, White, 1989), it is not surprising that concrete examples can facilitate comprehension and learning when instructional material is highly abstract.

There should be greater uncertainty regarding the benefits of two levels of description as the two levels approach each other along a scale of abstractness/generality. If two specific problems enable people to generate a more general description (Catrambone & Holyoak, 1989, Gick & Holyoak, 1983, Ross & Kennedy, 1990) then people who receive only specific problems should do as well as people who receive both levels of description. Or if the more general description enables people to generate a specific problem, then people who receive only a general description should do as well as people who receive both a general description and a specific problem. However, if it is not easy to move between levels of description, then providing both levels could be more helpful than providing only a single level. We evaluate these possibilities in Experiments 2

Table 3

Instructional Examples for Categorization Experiments

Equate Distances

An escaped convict walks away from a jail at an average speed of 4 mph. A tracker on horseback follows 2 hours later at an average speed of 7 mph. How long will it take the tracker to catch up with the convict?

Distance travelled by convict = Distance travelled by tracker

The Jones family drove to a national park at an average speed of 52 mph and returned along the same route at an average speed of 46 mph. How long did it take to reach the park if it took 1.5 hours longer to return?

Distance to the park = Distance from the park

Jack rides his motorcycle for 2 hours before his brother catches him after riding for 1.5 hours. What was Jack's speed if his brother's speed was 8 mph faster?

Distance travelled by Jack = Distance travelled by Jack's brother

Tim drove to his vacation home in 7 hours and returned the same route in 5 hours. How fast did he drive to his vacation home if his return speed was 18 mph faster?

Distance to vacation home = Distance from vacation home

Add Distances

Howard and Allan live 135 miles apart. They decide to drive toward each other, and Howard drives at 54 mph and Allan drives at 48 mph. How long will Allan drive before they meet if he leaves 1 hour after Howard?

Distance travelled by Howard + Distance travelled by Allan = Total Distance

Susan sailed 45 miles to her favorite island. She sailed 2.5 hours longer at 12 mph than she sailed at 9 mph. How long did she sail at 9 mph?

Distance travelled at 12 mph + Distance travelled at 9 mph = Total Distance

Two trains travel toward each other after leaving citites that are 750 miles apart. The first train travels for 7 hours and the second train travels for 4 hours before they meet. What was the speed of the first train if the second train is 8 mph faster?

Distance travelled by first train + Distance travelled by second train = Total Distance

An athlete trains by running for 1.5 hours and biking for 1 hour, covering a total distance of 25 miles. If his running speed is 10 mph slower than his biking speed, how fast does he run?

Distance travelled running + Distance travelled biking = Total Distance

Table 3 (continued)

Subtract Distances

Tom gave his brother Andy a 0.5 hour head start to see who could return home first. Andy beat his brother by 1 mile by jogging at 3 mph, compared to 4 mph for Tom. How long did it take Andy to reach home?

Distance travelled by Andy - Distance travelled by Tom = Difference in Distances

A man can walk at 1 mph after his accident, compared to 3 mph before his accident. He could walk 2 miles further before his accident by walking 0.5 hours longer at the faster speed. How long can he walk after his accident?

Distance before accident - Distance after accident = Difference in Distances

Karen's boat can travel 150 miles further in 8 hours than Jane's boat can travel in 10 hours. How fast is Karen's boat if it is 25 mph faster than Jane's boat?

Distance for Karen's boat - Distance for Jane's boat = Difference in Distances

A swimmer can swim 2 miles further when she swims for 3 hours with the current than when she swims for 4 hours against the current. If she can swim 1 mph faster with the current, how fast can she swim against the current?

Distance with current - Distance against current = Difference in Distances

Table 4

General Descriptions of Principles

Equate Distances

Two distances should be equated when one distance is equal to the other distance. For instance:

1. If two people travel the same route and one overtakes the other, then both travel the same distance. The correct equation shows that the distance travelled by one person equals the distance travelled by the other person.

2. If a person travels the same route on a round trip, then the distance travelled to the destination is the same as the distance travelled from the destination. The correct equation shows that the distance travelled to the destination equals the distance travelled from the destination.

Add Distances

Two travelled distances should be added when one distance plus the other distance combine to form the total distance. For instance:

1. If two people travel toward each other and meet, then the sum of the two travelled distances equals the total distance separating the two people. The correct equation shows that the distance travelled by one person plus the distance travelled by the other person equals the total distance they have to travel.

2. If a person travels at one speed and then changes speed, then the sum of the two travelled distances equals the total distance. The correct equation shows that the distance travelled at one speed plus the distance travelled at the other speed equals the total distance.

Subtract Distances

One travelled distance should be subtracted from the other when it is necessary to use the difference between the two distances. For instance:

1. If one person travels further than another, then the longer distance minus the shorter distance equals the difference in the two distances. The correct equation shows that the shorter distance should be subtracted from the longer distance if one distance is compared to another.

2. If a person travels different distances at different speeds, then the shorter distance should be subtracted from the longer distance if one distance is compared to another. The correct equation shows that the longer distance minus the shorter distance equals the difference in the distances.

and 3 by providing subjects with either specific problems, more general descriptions, or both to determine when two levels of description significantly improve performance over a single level of description.

Subjects in the *specific* instructional condition read two specific examples of each template. This material was identical to the information shown in Table 3 and included the equation that followed each problem. Subjects in the *general* instructional condition read a general description of each template. This material was identical to the information shown in Table 4. Subjects in the *combined* condition received both a general example and a specific example of each template. Each of the general descriptions was followed by a specific example -- either the first or second example from each category in Table 3.

The lack of a significant difference between the three instructional conditions is consistent with the interpretation that the difference between the two levels of description is insufficient to make a difference. Students who have two specific examples may be able to form a more general description of a template and students who have a general description may be reminded of a more specific example. If this interpretation is correct, then variations in the testing procedure should have little impact on the relative performance of the three groups because all three groups would have access to both levels of description.

An alternative hypothesis is that the potential advantage of receiving both levels of description was reduced by allowing students to refer back to the instructional material as they worked on the posttest. For instance, students who received only the general description had simultaneous access to both levels of description because the posttest contained specific examples. The test problems provided specific examples that could be use to instantiate the more general descriptions, reducing the informational difference between the general and the combined group.

Having access to the instructional material may also have benefitted students in the specific group. These students may not have formed a more general description, but nonetheless effectively used the specific examples because they could match problems on the posttest to the instructional examples. Research on how students use examples to solve physics problems showed that weaker students needed to reread the examples as they solved the test problems, suggesting that they were attempting to find a solution that could be copied (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). A limitation of this strategy is that information about the problems might be quickly forgotten, as documented by both Silver (1981) and Mayer (1982) for word problems. The usefulness of having only specific examples may therefore decline when students can not refer to the examples as they take the posttest.

The purpose of Experiment 3 was to determine whether receiving descriptions at two levels of generality would benefit students when they were not allowed to refer back to the instructional material. The procedure was identical to the procedure in Experiment 2 except students were instructed to not look at the instructional examples as they completed the posttest.

Subjects who received both the general and specific examples improved from 34% to 61% in identifying the correct operation. Subjects in the general condition improved from 33% to 43% correct, while subjects in the specific condition improved from 41% to 52% correct. Subjects in the combined condition showed significantly greater improvement than subjects in the general and the specific conditions, who did not differ significantly from each other.

The results revealed that two levels of description were significantly better than a single level when subjects were not allowed to refer back to the instructional material. This finding can not be explained solely by the amount of material presented during instruction. Subjects in the specific group had the same amount of material as subjects in the combined group but had two specific examples of each template rather than a specific example and a more general description. Subjects in the general group had less material, but the purpose of a general description is to reduce the reliance on specific examples by providing a summary of the essential characteristics of the examples.

Our findings do not support the hypothesis that subjects in the specific group formed their own general descriptions from the examples. If these subjects had generated such descriptions, their performance should be equivalent to subjects in the combined group. Providing general descriptions is therefore a useful supplement to specific examples when people have to rely on their memory for the examples. However, general descriptions that are not instantiated with specific examples may also have limited value, as illustrated by the modest improvement achieved by subjects in the general group.

We believe that the more general descriptions served two purposes in our study. First, they eliminated irrelevant story content that students might incorrectly use as a basis for selecting an analogous example. For instance, a frequent confusion that occurred in Experiment 1 involved a test problem that described a situation in which a police boat pursued another boat until it closed the gap to 5 miles. The most frequently selected analogy was an overtake problem in which a tracker caught up with an escaped convict. The more general description of an overtake problem does not mention a police story that likely caused the misclassification of the test problem.

A second benefit of a more general description is that it should help focus attention on relevant structural information. Specific problems contain structural information that is irrelevant to relating distances at the principle level. Formal relations also exist at lower (quantity) levels

such as specifying how much faster or how much more time one person travelled than another. The objective of the final two experiments was to determine whether these irrelevant relations can influence students' classifications (Reed, Zelmer, & Sanchez, 1991).

Effect of Speed Relations on Distance Classifications

In his book *Modularity of Mind* Jerry Fodor (1983) argued that the higher cognitive processes are difficult to study because there are so many sources of information that can influence thought. Others, such as Robert Sternberg (1985), have responded that higher cognitive processes become more modular with expertise because experts can quickly focus on the relevant information. One of the difficulties that students may have encountered in classifying the motion problems in Table 3 according the distance relations is that there are other irrelevant relations that may bias their classifications. Consider the problem:

Tim drove to his vacation home in 7 hours and returned the same route in 5 hours. How fast did he drive to his vacation home if his return speed was 18 mph faster? (Problem 5)

A particularly salient aspect of word problems may be the relations regarding the unknown variable expressed in the question. Reed, Zelmer, and Sanchez (1991) hypothesized that students would be biased toward a subtraction response if the question asked about an unknown value that was the slower of the two speeds, as in Problem 5. Conversely, students would be biased toward an addition response if the question asked about an unknown value that was the faster of the two speeds. This occurs if we change the question to:

2. How fast did he return from his vacation home if his return speed was 18 mph faster?

The question now asks about an unknown value (the return speed) that is the faster of the two speeds.

Another factor in our experimental design was whether the word *slower* or the word *faster* was used to describe the relation between the two speeds. We are told in Problem 5 that the return speed was 18 mph faster, but we could have been told that the initial speed was 18 mph slower:

3. How fast did he drive to his vacation home if his initial speed was 18 mph slower?

We hypothesized that using the word *slower* to describe the relation would bias subjects toward a subtraction response, and using the word *faster* to describe the relation would bias subjects toward an addition response. To complete the 2 (unknown) x 2 (relation) design we added a fourth question:

4. How fast did he return from his vacation home if his initial speed was 18 mph slower?

We gave 40 psychology students a classification task that consisted of 24 problems like those shown in Table 3, except that the unknown value was always speed. The correct answer was *equate distances* for 8 problems, *add distances* for 8 problems, and *subtract distances* for 8

problems. Each of the four variations of the question occurred for 2 of the 8 problems in these categories. Furthermore, there were four different test booklets so each problem would rotate through all four questions, as was illustrated for Tim's round trip.

Table 5 shows how the speed relations influenced the psychology students' classifications. We hypothesized that subjects would be biased toward responding that the two distances should be added when (1) the unknown value was the faster of the two speeds, and (2) the described relation between the speeds used the word *faster*. Both of these factors significantly influenced the addition responses. We also hypothesized that subjects would be biased toward responding that the two distances should be subtracted when (1) the unknown value was the slower of the two speeds, and (2) the described relation between the speeds used the word *slower*. The data in the second column show that both of these factors influenced the number of subtraction responses but only the unknown value had a significant effect. And finally, neither the unknown value nor the described relation influenced the number of equate responses.

Table 6 shows the effect of the distance relations on distance classifications. The first column shows that the greatest number of add responses occurred when the correct answer was *add* and the second column shows that the greatest number of subtract responses occurred when the correct answer was *subtract*. However, in neither of these two cases did the correct distance relation significantly influence the responses. Ironically, an irrelevant variable (speed) had a greater effect on these classifications than did the relevant variable (distance). Only for the unbiased equate category did the correct distance relation significantly affect classifications.

I previously mentioned that classification decisions should become more modular and less influenced by irrelevant relations as people acquire expertise. We therefore tested a second group of students who were majoring in mathematics and planned to teach mathematics at a secondary school. The lower half of Tables 5 and 6 show how the speed and distance relations influenced their classifications. The irrelevant speed relations had no effect on any of their responses, but the relevant distance relations significantly influenced all three of their responses. These students gave significantly more add responses when the correct answer was *add*, significantly more subtract responses when the correct answer was *subtract*, and significantly more equate responses when the correct answer was *equate*. It should be noted that this acquired modularity occurred before the classification task became very easy. Even the more expert group correctly classified only 2/3 of the problems, midway between chance and perfect performance.

Concluding Comments

I want to conclude by briefly discussing the generality of the schema-theory approach presented in this report. I introduced schema theory by referring to Thorndyke's description of a

Table 5

Effect of Speed Relations on Distance Classifications

Class	Speed		Responses		
	Unknown	Relation	Add	Subtract	Equate
Psychology	slower	slower	1.48	3.01	1.51
	slower	faster	1.88	2.66	1.46
	faster	slower	1.85	2.48	1.67
	faster	faster	2.20	2.21	1.59
Mathematics	slower	slower	1.70	1.95	2.35
	slower	faster	1.55	2.05	2.40
	faster	slower	1.80	1.95	2.25
	faster	faster	1.80	1.85	2.35

Table 6

Effect of Distance Relations on Distance Classifications

Class	Distance	Responses		
		Add	Subtract	Equate
Psychology	Add	2.70	3.53	1.77
	Subtract	2.55	3.85	1.60
	Equate	2.15	2.95	3.90
Mathematics	Add	5.10	1.55	1.35
	Subtract	1.25	4.85	1.90
	Equate	0.50	1.40	6.10

schema as a cluster of knowledge that provides a skeleton structure for a concept that can be instantiated by the detailed properties of a particular instance. The skeleton structure in our research is an equation and instantiation requires replacing general concepts (such as distance, rate, time) with quantities in a particular problem. Solving similar problems in the same domain requires modifying the equation because a similar problem has either more or less relevant quantities than the example. Solving isomorphic problems requires finding corresponding concepts across different domains. This may require representing problems at a higher level of abstraction in order to recognize the isomorphism.

I don't want to imply that all transfer conveniently fits into a schema-theory framework. Constructing equations for word problems seems to provide a nice fit but I am less enthusiastic about trying to fit my early research on the missionaries-cannibals problem (Reed, Ernst, & Banerji, 1974) into this approach. Ideas such as skeleton framework, instantiation of values, relevant versus irrelevant quantities, hierarchical levels that differ in generality, and procedural attachments have a more direct application in the word-problem domain than in the missionaries-cannibals domain.

A contrasting approach is a search-space framework in which subjects search for a solution (Newell & Simon, 1972). Gick (1986) has argued that a search is used when subjects do not have domain-specific knowledge for solving the problem and have to use general strategies such as means/end analysis. This is effective for small search spaces with a specific goal such as the missionaries/cannibals problem, but is not very helpful for algebra word problems where success is much more dependent on matching a test problem to domain-specific knowledge.

For very complex domains, expertise may depend on a combination of schema-based knowledge and general search strategies, as suggested by the study of electronic trouble shooting (Gott, Hall, Pokorny, Dibble, & Glaser, in press). Results revealed that skilled performers used three general types of schematic knowledge: device models of the equipment, general models of the trouble-shooting approach, and well-organized procedural knowledge for adapting to different equipment. The search approach characterized individual differences in breadth-first versus depth-first search. Good trouble shooters typically use a breadth-first search in which they first evaluate the major components before going into depth on any particular component.

In conclusion, we need to continue to look for effective ways of organizing knowledge to help students improve their transfer. In many ways the FERMI model (Larkin et al., 1988) is an ideal method of organization. A frame-like skeleton structure provides an organizing framework for attached procedures that can calculate needed values. Problems that are solved by common methods are organized together so students can see the commonality across different domains. For

complex domains such as statistics, physics, and algebra word problems, teachers, researchers, and textbook writers need to do more to make such an organizational framework apparent.

References

Aikins, J.S. (1983). Prototypical knowledge for expert systems. Artificial Intelligence, 20, 163-210.

Bobrow, D. G. & Winograd, T. (1977). An overview of KRL, a knowledge representation language. Cognitive Science, 1, 3-46.

Bransford, J.D., & Johnson, M.K. (1973). Consideration of some problems of comprehension. In W.G. Chase (Ed.), Visual information processing (pp. 383-438). Orlando, Florida: Academic Press.

Catrambone, R. & Holyoak, K.J. (1989). Overcoming contextual limitations on problem-solving transfer. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 1147-1156.

Catrambone, R., & Holyoak, K.J. (1990). Learning subgoals and methods for solving probability problems. Memory & Cognition, 18, 593-603.

Cheng, P.W., Holyoak, K.J., Nisbett, R.E., & Oliver, L.M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. Cognitive Psychology, 18, 293-328.

Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.

Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1, pp. 7-75). Hillsdale, NJ: Erlbaum.

Dellarosa, D. (1985). Abstraction of problem-type schemata through problem comparison (Tech Report No. 146). Boulder: University of Colorado, Institute of Cognitive Science.

Fodor, J. A. (1983). The modularity of mind. Cambridge, MA: The MIT Press.

Gick, M. (1986). Problem solving strategies. Educational Psychologist, 21, 99-120.

Gick, M. & Holyoak, K.J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gick, M. & Holyoak, K.J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Gott, S.P., Hall, E.P., Pokorny, R.A., Dibble, E., & Glaser, R. (1991). A naturalistic study of transfer: Adaptive expertise in technical domains. In D.K. Detterman & R.J. Sternberg (Eds.), Transfer on trial: Intelligence, cognition, and instruction. Norwood, NJ: Ablex Publishing Corp, in press.

Hinsley, D.A., Hayes, J.R., & Simon, H.A. (1977). From words to equations: Meaning and representation in algebra word problems. In P.A. Carpenter & M.A. Just (Eds.), Cognitive proceses in comprehension (pp. 89-106). Hillsdale, NJ: Erlbaum.

Holyoak, K.J. & Koh, K. (1987). Surface and structural similarity in analogical tranfer. *Memory & Cognition*, 15, 332-340.

Keiras, D. E. & Bovair, S. (1984). The role of a mental model in learning to operate a device. Cognitive Science, 8, 255-273.

Krutetskii, V.A. (1976). The psychology of mathematical abilities in school children. Chicago: The University of Chicago Press.

Larkin, J. H., Reif, F., Carbonell, J., & Guglietta, A. (1988). FERMI: A flexible expert reasoner with multi-domain inferencing. Cognitive Science, 12, 101-138.

Mayer, R.E. (1981). Frequency norms and structural analyses of algebra story problems into families, categories, and templates. Instructional Science, 10, 135-175.

Reed, S.K. (1987). A structure-mapping model for word problems. Journal of Experimental Psychology: Learning, Memory, & Cognition, 13, 124-139.

Reed, S.K. (1989). Constraints on the abstraction of solutions. Journal of Educational Psychology, 81, 532-540.

Reed, S. K. (1991). A schema-based theory of transfer. To appear in D.K. Detterman & R.J. Sternberg (Eds.), Transfer on trial: Intelligence, cognition, and instruction. Norwood, NJ: Ablex Publishing Corp.

Reed, S.K. Ackinclose, C.C., & Voss, A.A. (1990). Selecting analogous solutions: Similarity versus inclusiveness. Memory & Cognition, 18, 83-98.

Reed, S.K. & Bolstad, C.A. (1991). The use of examples and procedures in problem solving. Journal of Experimental Psychology: Learning, Memory, & Cognition, in press.

Reed, S.K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. Journal of Experimental Psychology: Learning, Memory, & Cognition, 11, 106-125.

Reed, S.K., Ernst, G.W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. Cognitive Psychology, 6, 436-450.

Reed, S.K. & Ettinger, M. (1987). Usefulness of tables for solving word problems. Cognition and Instruction, 4, 43-59.

Reed, S.K. & Sanchez, F.J. (1990). Generality of descriptions in categorizing algebra word problems. Manuscript submitted for publication.

Reed, S.K., Zelmer, R., & Sanchez, F.J. (1991). Acquisition of modularity across relational assignments in word problems. Manuscript in preparation.

Rosch, E., Mervis, C.B., Gray, W.D. Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 7, 573-605.

Ross, B.H. (1989). Remindings in learning and instruction. In S. Vosniadou and A. Ortony (Eds.), Similarity and analogical reasoning (pp. 438-469). Cambridge: Cambridge University Press.

Ross, B.H. & Kennedy, P.T. (1990). Generalizing from the use of earlier examples in problem solving. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 42-55.

Schoenfeld, A.H. & Hermann, D.J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. Journal of Experimental Psychology: Learning, Memory, & Cognition, 5, 484-494.

Silver, E.A. (1981). Recall of mathematical problem information: Solving related problems. Journal for Research in Mathematics Education, 12, 54-64.

Smith, E. E., & Goodman, L. (184). Understanding written instructions: The role of explanatory schema. Cognition and Instruction, 1, 359-396.

Sternberg, R.J. (1985). Controlled versus automatic processing. Behavioral and Brain Sciences, 8, 32-33.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. Cognition and Instruction, 2, 59-89.

Thorndyke, P.W. (1984). Applications of schema theory in cognitive research. In J.R. Anderson & S.M. Kosslyn (Eds.), Tutorials in learning and memory (pp. 167-191). San Francisco: Freeman.

Viscuso, S. R., & Spoehr, K. T. (1986). How does a mental model facilitate comprehension of instructions? Paper presented at the 27th Annual Meeting of the Psychonomic Society, November, 1986.

White, B.Y. (1989). The role of intermediate abstractions in understanding science and mathematics. Proceedings of the Eleventh Annual Meeting of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.

Winston, P.H. (1980). Learning and reasoning by analogy. Communications of the ACM, 23, 689-703.

Publications

Reed, S.K. (1989). Constraints on the abstraction of solutions. Journal of Educational Psychology, 81, 532-540.

Reed, S.K. Ackinclose, C.C., & Voss, A.A. (1990). Selecting analogous solutions: Similarity versus inclusiveness. Memory & Cognition, 18, 83-98.

Reed, S.K. & Sanchez, F.J. (1990). Generality of descriptions in categorizing algebra word problems. Manuscript submitted for publication.

Reed, S. K. (1991). A schema-based theory of transfer. To appear in D.K. Detterman & R.J. Sternberg (Eds.), Transfer on trial: Intelligence, cognition, and instruction. Norwood, NJ: Ablex Publishing Corp.

Reed, S.K. (1991). Review of S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning. American Journal of Psychology, in press.

Reed, S.K. & Bolstad, C.A. (1991). The use of examples and procedures in problem solving. Journal of Experimental Psychology: Learning, Memory, & Cognition, in press.

Reed, S.K., Zelmer, R., & Sanchez, F.J. (1991). Acquisition of modularity across relational assignments in word problems. Manuscript in preparation.

Presentations

Conferences

Paper presentation, Psychonomics Society, Chicago, November, 1988.

Progress report, Air Force Office of Scientific Research, Arlington, VA, November, 1989.

Symposium presentation, American Educational Research Association, Boston, April, 1990.

Paper presentation (planned), Psychonomics Society, San Francisco, November, 1991.

Colloquia

I reported on various aspects of this research at colloquia presented at the University of South Florida, Florida International University, San Diego State University, UCLA, the Georgia Institute of Technology, Claremont Graduate School, and the University of California, Santa Cruz.

Personnel

Cheryl Actor (Bolstad), Supervised M.A. Thesis (Combining Examples and Procedures, May, 1988) at Florida Atlantic University.

Audrey Voss, Graduate student in M.A. program, Florida Atlantic University.

Javier Sanchez, Graduate student in M.A. program, San Diego State University.

Robert Zelmer, Post-B.A. student, San Diego State University.